



PCT

## RAW SEQUENCE LISTING

DATE: 09/09/2004

PATENT APPLICATION: US/10/506,548

TIME: 15:53:01

Input Set : A:\Q83405 Sequence Listing.txt  
 Output Set: N:\CRF4\09092004\J506548.raw

4 <110> APPLICANT: SEIKAGAKU CORPORATION  
 5 AMERSHAM BIOSCIENCES K.K.  
 6 NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY  
 7 Hisashi NARIMATSU  
 8 Shigemi SUGIOKA  
 9 Hideo MOCHIZUKI  
 11 <120> TITLE OF INVENTION: Sulfotransferase, peptide thereof and DNA encoding the same  
 13 <130> FILE REFERENCE: Q83405  
 C--> 15 <140> CURRENT APPLICATION NUMBER: US/10/506,548  
 C--> 15 <141> CURRENT FILING DATE: 2004-09-03  
 15 <150> PRIOR APPLICATION NUMBER: PCT/JP03/02500  
 16 <151> PRIOR FILING DATE: 2003-03-04  
 18 <150> PRIOR APPLICATION NUMBER: JP 2002-57527  
 19 <151> PRIOR FILING DATE: 2002-03-04  
 21 <150> PRIOR APPLICATION NUMBER: JP 2002-245994  
 22 <151> PRIOR FILING DATE: 2002-08-26  
 24 <160> NUMBER OF SEQ ID NOS: 16  
 26 <210> SEQ ID NO: 1  
 27 <211> LENGTH: 1041  
 28 <212> TYPE: DNA  
 29 <213> ORGANISM: Homo sapiens  
 31 <220> FEATURE:  
 32 <221> NAME/KEY: CDS  
 33 <222> LOCATION: (1)..(1041)  
 35 <400> SEQUENCE: 1  
 36 atg cta ttc aaa cag cag gcg tgg ctg aga cag aag ctc ctg gtg ctg 48  
 37 Met Leu Phe Lys Gln Gln Ala Trp Leu Arg Gln Lys Leu Leu Val Leu  
 38 1 5 10 15  
 40 gga agc ctt gcc gtt ggg agt ctc ctg tat cta gtc gcc aga gtt ggg 96  
 41 Gly Ser Leu Ala Val Gly Ser Leu Leu Tyr Leu Val Ala Arg Val Gly  
 42 20 25 30  
 44 agc ttg gat agg cta caa ccc att tgc ccc att gaa ggt cga ctg ggt 144  
 45 Ser Leu Asp Arg Leu Gln Pro Ile Cys Pro Ile Glu Gly Arg Leu Gly  
 46 35 40 45  
 48 gga gcc cgc act cag gct gaa ttc cca ctt cgc gcc ctg cag ttt aag 192  
 49 Gly Ala Arg Thr Gln Ala Glu Phe Pro Leu Arg Ala Leu Gln Phe Lys  
 50 50 55 60  
 52 cgt ggc ctg ctg cac gag ttc cgg aag ggc aac gct tcc aag gag cag 240  
 53 Arg Gly Leu Leu His Glu Phe Arg Lys Gly Asn Ala Ser Lys Glu Gln  
 54 65 70 75 80  
 56 gtt cgc ctc cat gac ctg gtc cag cag ctc ccc aag gcc att atc att 288  
 57 Val Arg Leu His Asp Leu Val Gln Gln Leu Pro Lys Ala Ile Ile Ile  
 58 85 90 95

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/506,548

DATE: 09/09/2004  
TIME: 15:53:01

Input Set : A:\Q83405 Sequence Listing.txt  
Output Set: N:\CRF4\09092004\J506548.raw

60	ggg	gtg	agg	aaa	gga	ggc	aca	agg	gcc	ctg	ctt	gaa	atg	ctg	aac	cta	336	
61	Gly	Val	Arg	Lys	Gly	Gly	Thr	Arg	Ala	Leu	Leu	Glu	Met	Leu	Asn	Leu		
62									100		105				110			
64	cat	ccg	gca	gta	gtc	aaa	gcc	tct	caa	gaa	atc	cac	ttt	ttt	gat	aat	384	
65	His	Pro	Ala	Val	Val	Lys	Ala	Ser	Gln	Glu	Ile	His	Phe	Phe	Asp	Asn		
66									115		120				125			
68	gat	gag	aat	tat	ggt	aag	ggc	att	gag	tgg	tat	agg	aaa	aag	atg	cct	432	
69	Asp	Glu	Asn	Tyr	Gly	Lys	Gly	Ile	Glu	Trp	Tyr	Arg	Lys	Lys	Met	Pro		
70									130		135				140			
72	ttt	tcc	tac	cct	cag	caa	atc	aca	att	gaa	aag	agc	cca	gca	tat	ttt	480	
73	Phe	Ser	Tyr	Pro	Gln	Gln	Ile	Thr	Ile	Glu	Lys	Ser	Pro	Ala	Tyr	Phe		
74	145						150			155					160			
76	atc	aca	gag	gag	gtt	cca	gaa	agg	att	tac	aaa	atg	aac	tca	tcc	atc	528	
77	Ile	Thr	Glu	Glu	Val	Pro	Glu	Arg	Ile	Tyr	Lys	Met	Asn	Ser	Ser	Ile		
78							165			170					175			
80	aag	ttg	ttg	atc	att	gtc	agg	gag	cca	acc	aca	aga	gct	att	tct	gat	576	
81	Lys	Leu	Leu	Ile	Ile	Val	Arg	Glu	Pro	Thr	Thr	Arg	Ala	Ile	Ser	Asp		
82							180			185					190			
84	tat	act	cag	gtg	cta	gag	ggg	aag	gag	agg	aag	aac	aaa	act	tat	tac	624	
85	Tyr	Thr	Gln	Val	Leu	Glu	Gly	Lys	Glu	Arg	Lys	Asn	Lys	Thr	Tyr	Tyr		
86							195			200					205			
88	aag	ttt	gag	aag	ctg	gcc	ata	gac	cct	aat	aca	tgc	gaa	gtg	aac	aca	672	
89	Lys	Phe	Glu	Lys	Leu	Ala	Ile	Asp	Pro	Asn	Thr	Cys	Glu	Val	Asn	Thr		
90							210			215					220			
92	aaa	tac	aaa	gca	gta	aga	acc	agc	atc	tac	acc	aaa	cat	ctg	gaa	agg	720	
93	Lys	Tyr	Lys	Ala	Val	Arg	Thr	Ser	Ile	Tyr	Thr	Lys	His	Leu	Glu	Arg		
94	225						230			235					240			
96	tgg	ttg	aaa	tac	ttt	cca	att	gag	caa	ttt	cat	gtc	gtc	gat	gga	gat	768	
97	Trp	Leu	Lys	Tyr	Phe	Pro	Ile	Glu	Gln	Phe	His	Val	Val	Asp	Gly	Asp		
98							245			250					255			
100	cgc	ctc	atc	acg	gaa	cct	ctg	cca	gaa	ctt	cag	ctc	gtg	gag	aag	ttc	816	
101	Arg	Leu	Ile	Thr	Glu	Pro	Leu	Pro	Glu	Leu	Gln	Leu	Val	Glu	Lys	Phe		
102							260			265					270			
104	cta	aat	ctg	cct	cca	agg	ata	agt	caa	tac	aat	tta	tac	tcc	aat	gct	864	
105	Leu	Asn	Leu	Pro	Pro	Arg	Ile	Ser	Gln	Tyr	Asn	Leu	Tyr	Phe	Asn	Ala		
106							275			280					285			
108	acc	aga	ggg	ttt	tac	tgc	ttg	cgg	ttt	aat	att	atc	ttt	aat	aag	tgc	912	
109	Thr	Arg	Gly	Phe	Tyr	Cys	Leu	Arg	Phe	Asn	Ile	Ile	Phe	Asn	Lys	Cys		
110							290			295					300			
112	ctg	gcg	ggc	agc	aag	ggg	cgc	att	cat	cca	gag	gtg	gac	ccc	tct	gtc	960	
113	Leu	Ala	Gly	Ser	Lys	Gly	Arg	Ile	His	Pro	Glu	Val	Asp	Pro	Ser	Val		
114	305						310			315					320			
116	att	act	aaa	ttg	cgc	aaa	ttc	ttt	cat	cct	ttt	aat	caa	aaa	ttt	tac	1008	
117	Ile	Thr	Lys	Leu	Arg	Lys	Phe	Phe	His	Pro	Phe	Asn	Gln	Lys	Phe	Tyr		
118							325			330					335			
120	cag	atc	act	ggg	agg	aca	ttg	aac	tgg	ccc	taa						1041	
121	Gln	Ile	Thr	Gly	Arg	Thr	Leu	Asn	Trp	Pro								
122							340			345								
125	<210>	SEQ	ID	NO:	2													

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/506,548

DATE: 09/09/2004  
TIME: 15:53:01

Input Set : A:\Q83405 Sequence Listing.txt  
Output Set: N:\CRF4\09092004\J506548.raw

126 <211> LENGTH: 346  
127 <212> TYPE: PRT  
128 <213> ORGANISM: Homo sapiens  
130 <400> SEQUENCE: 2  
131 Met Leu Phe Lys Gln Gln Ala Trp Leu Arg Gln Lys Leu Leu Val Leu  
132 1 5 10 15  
133 Gly Ser Leu Ala Val Gly Ser Leu Leu Tyr Leu Val Ala Arg Val Gly  
134 20 25 30  
135 Ser Leu Asp Arg Leu Gln Pro Ile Cys Pro Ile Glu Gly Arg Leu Gly  
136 35 40 45  
137 Gly Ala Arg Thr Gln Ala Glu Phe Pro Leu Arg Ala Leu Gln Phe Lys  
138 50 55 60  
139 Arg Gly Leu Leu His Glu Phe Arg Lys Gly Asn Ala Ser Lys Glu Gln  
140 65 70 75 80  
141 Val Arg Leu His Asp Leu Val Gln Gln Leu Pro Lys Ala Ile Ile Ile  
142 85 90 95  
143 Gly Val Arg Lys Gly Gly Thr Arg Ala Leu Leu Glu Met Leu Asn Leu  
144 100 105 110  
145 His Pro Ala Val Val Lys Ala Ser Gln Glu Ile His Phe Phe Asp Asn  
146 115 120 125  
147 Asp Glu Asn Tyr Gly Lys Gly Ile Glu Trp Tyr Arg Lys Lys Met Pro  
148 130 135 140  
149 Phe Ser Tyr Pro Gln Gln Ile Thr Ile Glu Lys Ser Pro Ala Tyr Phe  
150 145 150 155 160  
151 Ile Thr Glu Glu Val Pro Glu Arg Ile Tyr Lys Met Asn Ser Ser Ile  
152 165 170 175  
153 Lys Leu Leu Ile Ile Val Arg Glu Pro Thr Thr Arg Ala Ile Ser Asp  
154 180 185 190  
155 Tyr Thr Gln Val Leu Glu Gly Lys Glu Arg Lys Asn Lys Thr Tyr Tyr  
156 195 200 205  
157 Lys Phe Glu Lys Leu Ala Ile Asp Pro Asn Thr Cys Glu Val Asn Thr  
158 210 215 220  
159 Lys Tyr Lys Ala Val Arg Thr Ser Ile Tyr Thr Lys His Leu Glu Arg  
160 225 230 235 240  
161 Trp Leu Lys Tyr Phe Pro Ile Glu Gln Phe His Val Val Asp Gly Asp  
162 245 250 255  
163 Arg Leu Ile Thr Glu Pro Leu Pro Glu Leu Gln Leu Val Glu Lys Phe  
164 260 265 270  
165 Leu Asn Leu Pro Pro Arg Ile Ser Gln Tyr Asn Leu Tyr Phe Asn Ala  
166 275 280 285  
167 Thr Arg Gly Phe Tyr Cys Leu Arg Phe Asn Ile Ile Phe Asn Lys Cys  
168 290 295 300  
169 Leu Ala Gly Ser Lys Gly Arg Ile His Pro Glu Val Asp Pro Ser Val  
170 305 310 315 320  
171 Ile Thr Lys Leu Arg Lys Phe Phe His Pro Phe Asn Gln Lys Phe Tyr  
172 325 330 335  
173 Gln Ile Thr Gly Arg Thr Leu Asn Trp Pro  
174 340 345  
177 <210> SEQ ID NO: 3

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/506,548

DATE: 09/09/2004

TIME: 15:53:01

Input Set : A:\Q83405 Sequence Listing.txt  
 Output Set: N:\CRF4\09092004\J506548.raw

178 <211> LENGTH: 12  
 179 <212> TYPE: DNA  
 180 <213> ORGANISM: Artificial Sequence  
 182 <220> FEATURE:  
 183 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' Primer for PCR  
 185 <400> SEQUENCE: 3  
 186 ctacaaccca tt 12  
 189 <210> SEQ ID NO: 4  
 190 <211> LENGTH: 12  
 191 <212> TYPE: DNA  
 192 <213> ORGANISM: Artificial Sequence  
 194 <220> FEATURE:  
 195 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' Primer for PCR  
 197 <400> SEQUENCE: 4  
 198 ttagggccag tt 12  
 201 <210> SEQ ID NO: 5  
 202 <211> LENGTH: 12  
 203 <212> TYPE: DNA  
 204 <213> ORGANISM: Artificial Sequence  
 206 <220> FEATURE:  
 207 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' Primer for PCR  
 209 <400> SEQUENCE: 5  
 210 atgctattca aa 12  
 213 <210> SEQ ID NO: 6  
 214 <211> LENGTH: 24  
 215 <212> TYPE: DNA  
 216 <213> ORGANISM: Artificial Sequence  
 218 <220> FEATURE:  
 219 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' Primer for PCR (GP-  
 226)  
 221 <400> SEQUENCE: 6  
 222 cggaaactcggt gcagcaggcc acgc 24  
 225 <210> SEQ ID NO: 7  
 226 <211> LENGTH: 24  
 227 <212> TYPE: DNA  
 228 <213> ORGANISM: Artificial Sequence  
 230 <220> FEATURE:  
 231 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR (GP-  
 224)  
 233 <400> SEQUENCE: 7  
 234 tcgacacctca atggggcaaa tggg 24  
 237 <210> SEQ ID NO: 8  
 238 <211> LENGTH: 25  
 239 <212> TYPE: DNA  
 240 <213> ORGANISM: Artificial Sequence  
 242 <220> FEATURE:  
 243 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR  
 (SFTex2F)  
 245 <400> SEQUENCE: 8  
 246 actggggaaac cagaaaaatg aaaag 25  
 249 <210> SEQ ID NO: 9  
 250 <211> LENGTH: 25

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/506,548

DATE: 09/09/2004  
TIME: 15:53:01

Input Set : A:\Q83405 Sequence Listing.txt  
Output Set: N:\CRF4\09092004\J506548.raw

251 <212> TYPE: DNA  
252 <213> ORGANISM: Artificial Sequence  
254 <220> FEATURE:  
255 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' primer for PCR  
(SFTex2R)  
257 <400> SEQUENCE: 9  
258 gtgtctccag gcacaacaca tagtg 25  
261 <210> SEQ ID NO: 10  
262 <211> LENGTH: 55  
263 <212> TYPE: DNA  
264 <213> ORGANISM: Artificial Sequence  
266 <220> FEATURE:  
267 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR  
(SFTgateF2)  
269 <400> SEQUENCE: 10  
270 ggggacaagt ttgtacaaaa aaggcaggctt ctttaagcgt ggcctgctgc acgag 55  
273 <210> SEQ ID NO: 11  
274 <211> LENGTH: 53  
275 <212> TYPE: DNA  
276 <213> ORGANISM: Artificial Sequence  
278 <220> FEATURE:  
279 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' primer for PCR  
(SFTgateTstop)  
281 <400> SEQUENCE: 11  
282 ggggaccact ttgtacaaga aagctgggtt tagggccagt tcaatgtcct ccc 53  
285 <210> SEQ ID NO: 12  
286 <211> LENGTH: 22  
287 <212> TYPE: PRT  
288 <213> ORGANISM: Artificial Sequence  
290 <220> FEATURE:  
291 <223> OTHER INFORMATION: Description of Artificial Sequence: Ig kappa signal sequence  
293 <400> SEQUENCE: 12  
294 Met His Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala Ser  
295 1 5 10 15  
297 Val Ile Met Ser Arg Gly  
298 20  
301 <210> SEQ ID NO: 13  
302 <211> LENGTH: 7  
303 <212> TYPE: PRT  
304 <213> ORGANISM: Artificial Sequence  
306 <220> FEATURE:  
307 <223> OTHER INFORMATION: Description of Artificial Sequence: FLAG peptide  
309 <400> SEQUENCE: 13  
310 Asp Tyr Lys Asp Asp Asp Lys  
311 1 5  
314 <210> SEQ ID NO: 14  
315 <211> LENGTH: 94  
316 <212> TYPE: DNA  
317 <213> ORGANISM: Artificial Sequence  
319 <220> FEATURE:  
320 <223> OTHER INFORMATION: Description of Artificial Sequence: OT3 seugnce  
322 <400> SEQUENCE: 14

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/506,548

DATE: 09/09/2004

TIME: 15:53:02

Input Set : A:\Q83405 Sequence Listing.txt  
Output Set: N:\CRF4\09092004\J506548.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No  
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date